Complete

H3.2. Here we first prove that, f is μ-strongly convex $\Longleftrightarrow$ f^* is $\frac{1}{\mu}$-smooth.
Then according to $f^{**} = f$, f is L-smooth $\Longleftrightarrow$ f^* is $\frac{1}{L}$-strongly convex.
Finally for the $\frac{1}{L}$-strongly convex and $\frac{1}{\mu}$-smooth f^*, prove the convergence of dual ascent.

① Prove f is μ-strongly convex $\Longleftrightarrow$ f^* is $\frac{1}{\mu}$-smooth.

Proof: Let $\mathcal{L} = f - y^T x$, so $\mathcal{L}$ is also μ-strongly convex.

of "$\Rightarrow$": Defining $x_u = \nabla f^*(u)$, $x_v = \nabla f^*(v)$. For a minimizer x^*, there are:

$$\mathcal{L}(x) \geq \mathcal{L}(x^*) + \frac{\mu}{2}\|x - x^*\|_2^2 \quad \text{since } \nabla\mathcal{L}(x^*) = 0.$$

So:
$$\begin{cases} f(x_v) - u^T x_v \geq f(x_u) - u^T x_u + \frac{\mu}{2}\|x_u - x_v\|_2^2 \\ f(x_u) - v^T x_u \geq f(x_v) - v^T x_v + \frac{\mu}{2}\|x_u - x_v\|_2^2 \end{cases}$$

Adding them together: $(u^T - v^T)(x_u - x_v) \geq \mu\|x_u - x_v\|_2^2$

With Cauchy-Schwartz: $(u^T - v^T)(x_u - x_v) \leq \|u^T - v^T\|\,\|x_u - x_v\|$

So $\|x_u - x_v\| \leq \|u - v\|/\mu$

$\Rightarrow$ ∇f^* is Lipschitz with $\frac{1}{\mu}$ $\Rightarrow$ f^* is $\frac{1}{\mu}$-smooth.

Proof of "$\Leftarrow$": For $\frac{1}{\mu}$-smooth f^*, ~~there is: let $g_x(\eta) = f(\eta) - \nabla g(x)^T \eta$.~~

let $g_x(\eta) = f^*(\eta) - \nabla f^*(x)^T \eta$.

$$\text{so} \quad g_x(\eta) \leq g_x(y) + \nabla g_x(y)^T(\eta - y) + \frac{1}{2\mu}\|\eta - y\|_2^2.$$

Minimizing each side over η, and rearranging, there is:

$$\frac{\mu}{2}\|\nabla f^*(x) - \nabla f^*(y)\|_2^2 \leq f^*(y) - f^*(x) + \nabla f^*(x)^T(x - y).$$

Exchange x and y: $\frac{\mu}{2}\|\nabla f^*(x) - \nabla f^*(y)\|_2^2 \leq f^*(x) - f^*(y) + \nabla f^*(y)^T(y - x)$

Adding together: $\mu\|\nabla f^*(x) - \nabla f^*(y)\|_2^2 \leq [\nabla f^*(x) - \nabla f^*(y)]^T(x - y)$

Let $u = \nabla f^*(x)$, $v = \nabla f^*(y)$, so $(x - y)^T(u - v) \geq \mu\|u - v\|_2^2$

With the proof of "$\Leftarrow$" and "$\Rightarrow$", Q.E.D.

② Study the convergence of dual ascent.

Dual ascent is another G.D with different params

For $\frac{1}{\mu}$-smooth f^*, a fixed step $\alpha = 1/\frac{1}{\mu} = \mu$ is often selected.

$$x_{k+1} = x_k + \mu \nabla f^*(x_k)$$

$\because$ $f^*(y) \leq f^*(x) + \nabla f^*(x)(y-x) + \frac{1}{2\mu}\|y-x\|_2^2$

$$f^*(x_{k+1}) \leq f^*(x_k) + \frac{3\mu}{2}\|\nabla f^*(x_k)\|_2^2 \quad ①$$

$$\Rightarrow f^*(x_k) - f^*(x^*) \leq \frac{3\mu}{2k}\|x_0 - x^*\|_2^2$$

to guarantee $f^*(x_k) - f^*(x^*) \leq \epsilon$, there should be $k \geq \frac{3\mu}{2\epsilon}\|x_0 - x^*\|_2^2$

$\therefore$ $k \sim O(\frac{1}{\epsilon})$ sublinear.

For $\frac{1}{L}$-strongly convex f^*, there are:

$$f^*(y) \geq f^*(x) + \nabla f^*(x)(y-x) + \frac{1}{2L}\|y-x\|_2^2.$$

Minimize this lower bound by taking the gradient w.r.t y and setting it to 0:

$$\nabla f^*(x) + \frac{1}{L}(y-x) = 0.$$

So the lower bound is: $f^*(y) \geq f^*(x) - L\|\nabla f^*(x)\|_2^2 + \frac{L}{2}\|\nabla f^*(x)\|_2^2$

$$= f^*(x) - \frac{L}{2}\|\nabla f^*(x)\|_2^2$$

$$\therefore \|\nabla f^*(x)\|_2^2 \geq \frac{2}{L}[f^*(x) - f^*(y)]$$

let $y = x^*$ and combined with ①:

$$f^*(x_{k+1}) - f^*(x^*) \leq f^*(x_k) - f^*(x^*) + \frac{3\mu}{L}[f^*(x_k) - f^*(x^*)]$$

$$= (1 + \frac{3\mu}{L})[f^*(x_k) - f^*(x^*)].$$

$$\therefore f^*(x_k) - f^*(x^*) \leq (1 + \frac{3\mu}{L})^k \cdot [f^*(x_0) - f^*(x^*)] \leq \epsilon$$

$$\Rightarrow k \geq \frac{\log[(f^*(x_0) - f^*(x^*))/\epsilon]}{\log[(L+3\mu)/L]} \sim O(\log\frac{1}{\epsilon})$$

linear.

However, the solution is not guaranteed to be primal feasible, since while the dual variables converge to a solution, the primal variables may not satisfy the constraints of the original problem.